

統計學概要

附普考與特考試題詳解

高東正 編著

統計學概要

附普考與特考試題詳解

高東正 編著

科技圖書股份有限公司

版 權 所 有
翻 印 必 究

統計學概要

編著者：高 東 正

總經銷：科技圖書股份有限公司

台北市博愛路185號二樓

電話：3 1 1 0 9 5 3

郵政劃撥 1 5 6 9 7 號

六十八年五月初版

特價新台幣125元

序

歷史的軌跡明確地昭示：人類自有政府組織以來，就有著統計之意識與需要，而今日科技的發達，管理的講究，更使統計方法顯居要位。公家機關、私人機構的招錄，往往需考統計學，若能擁有一本系統之書，當能掌握準備之方向，以收事半功倍之效。

依筆者之體驗，高考與普考於程度上有著不可否認的差距，基於街坊間不乏缺欠系統之作，且籠統地將高考與普考混為一談，並附上研究所及各校之試題，此類書籍恐不適參加普考者研讀，因而著手編著此書。於其中捨棄高考及較堅深之理論，以普考、特考及一般機關考試之題目，衍敘切用正確的統計觀念與方法，以供有志參加普考之上進青年，於準備之際有一正確的方向與良伴，並供統計系之初學者，奠定不惑的觀念。

本書共分十二章，每章依下列次序編寫：

(一)導言：以扼要之方式，使讀者對該章之旨意略有所識。

(二)要義：以簡潔之筆調，完整地定義該章之名詞及定理，可能用到之數學重要式子，則以備用公式之型態出現。

(三)普考試題：網羅 67 年以前之歷屆考題，編寫正確、簡潔的解法。

基於試題往往是綜合數章而成，因而依下列兩原則編之。

①解釋名詞之題目，依出現之章節次序編入。

②計算題目視題意之範圍編入適當的章節。

(四)習題範例：甄選重要之題目以及專科教材之要題以資補充，並詳證重要公式、定理及性質。

本書的編成，承科技圖書公司及東府出版社的支持，在校師長的教誨以及蔡秋蘭、黃素貞、洪玉琴、張顏忠、尹守誠、許重榮等學友的協助，在此一併誌謝。

本書若有漏誤，尚請師長賜教包涵為祈。

願本書與有志者共同奮鬥，並預祝名題金榜

編著 高 東 正

謹識於分子尾

目 錄

第一章	統計圖表與次數分配	1
第二章	平均數、中位數、衆數與其他中央趨勢值的測量	42
第三章	差異數量	80
第四章	機 率 論	119
第五章	隨機變數及其分佈	138
第六章	統計量分佈	185
第七章	偏態、峯度與動差	217
第八章	統計推估	247
第九章	假設檢定	292
第十章	簡單迴歸與相關	366
第十一章	指 數	413
第十二章	時間數列	438

第一章

統計圖表與次數分配

*****導 言*****

本章擬分兩節，首先描述有關統計圖表的製作過程；第二節再由統計圖表發展至次數分配。

一、統計學

乃搜集、整理、表現和分析資料，並根據分析結果導出適當之結論、作合理決策的科學方法。現代統計學家對此名詞的引用，常有三種不同的含義：(一)指統計資料；(二)指統計方法；(三)指統計原理。然而它係一種解決問題的工具或手段，故很多學者直稱之為統計方法。

二、統計資料

藉實驗或調查而得有關社會現象或自然現象之群體特性的數字。

種類：

1 原始資料：統計人員為了某種目的，經由實地調查或逕由資料來源處收集，或其他機構搜集而來，經任何簡化與整理的數據稱之。

2 次級資料：原始資料經整理、分析、簡化過程所得之結果以供利用的數據稱之。

3 靜態資料：表示某現象在某一特定時刻之靜止狀態資料。強調瞬間標準時刻，大多由調查而得。

4 動態資料：表示某現象於某特定時期內繼續發生之演變過程的資料；大多由登記而得。

2 統計學概論

性 質：

- 1 具數字性：統計資料係由數計或度量而得，因而均是數字的資料。
- 2 具群體性：統計資料係大量地觀察群體而得。
- 3 具客觀性：一切統計資料均須由實際調查、實驗或登記而來，不可由主觀之臆測或估計而得。

要 素：

- 1 時間。
- 2 空間。
- 3 特性。

搜集方法：

- 1 調查：依方式分：

- (A) 親自調查。
- (B) 委託調查。
- (C) 通訊調查。

依範圍分：

- (A) 普查：將所欲研究的某種現象之全部個體，一一加以調查。
- (B) 抽樣調查：於所欲研究的某種現象之全部個體中，選取一部份而加以調查。

- 2 登記。
- 3 實驗。
- 4 次級資料的收集。

整理步驟：

- 1 分類：將事物依其特性，把相同或相似者合為一類：
一般標準如下：

- (A) 時間標準：以事物發生之時間先後而加以分類。
- (B) 地域標準：以事物發生所在之區域而加以分類。
- (C) 特性標準：依事物之質特性或量特性加以分類。

原 則：

- (A) 互斥性：所分之各類必須互相排斥，凡能歸入某類者，絕不能再歸入他類，此目的在於分類時避免重複。
- (B) 周延性：所分之各類必須所包括所應分類的一切事物。

2 歸類：將原始資料分別歸入其所應屬類別的過程。

主要方法：

(A)人工處理。

(B)電腦處理。(詳見第1-8題)

3 列表：將次級資料依所研究之目的做成統計表或統計圖，以顯示資料的資訊。

三、統計圖

乃為一種平面或立體之圖形，係利用點之多寡、線之長短、面積或體積之大小、顏色之濃淡、線條之疏密或曲線之傾斜角度，以表示統計表內各種數字的大小程度、變動情形、分佈狀況或相互關係的一種圖表。

種類：詳見第1-22題。

製圖步驟：詳見第1-11題。

結構之細節：詳見第1-23題。

目的：

1 使閱者花費甚少的時間，即可對某一事實獲得明確、具體的概念。

2 引起閱者對該事實的興趣，避免冗長的文字、數字說明。

3 可以顯示數量間的相互關係。

4 便於多種複雜現象的比較，可供研究者分析及說明，又可供商業活動、廣告宣傳及機關工作表現績效上使用。

四、統計方法

將複雜的統計資料，運用統計的知識使之簡單化、系統化與表徵化，以探求群體之特質。

步驟：

1 搜集資料。

2 整理資料。

3 分析現況。

4 推測未來。

應用範圍：

1 經濟統計——應用於經濟學上。

4 統計學概論

- 2 人口統計——應用於人口問題之研究。
- 3 生物統計——應用於農作物之改良、牲畜之育種以及遺傳之研究。
- 4 農業統計——應用於農村調查、農作經濟、農情預測以及農情報告等之研究問題。
- 5 醫學統計——應用於細菌擴散、病理分析等之研究問題。
- 6 實驗設計——應用於農作物之田間或工業上試驗之事先設計以及事後分析。
- 7 品質管制——應用於工業控制產品品質之問題。
- 8 其他——如應用於心理、教育等心理統計及教育統計。

五、有效數字

表示一數值之精確程度的數字稱之。

判定原則：

- 1 經簡化後非零之位數均為有效數字。
- 2 零居於兩有效數字之間或零居於有效數字之後，但含小數點時，恒為有效數字。
- 3 以零開頭之數字，其居於第一個非零數字前之零不為有效數字。
- 4 零居於最後一個有效數字之後，則有效位數之個數不定。

六、誤 差

- 1 絕對誤差：近似值與真值之差。
- 2 可能誤差：乃統計測量值數字之最末一位數的半個單位。
例：某人身高 142 公分，則其實際身高可能在 142 ± 0.5 公分之間，此 0.5 公分即是可能誤差。
- 3 相對誤差：乃絕對誤差與真值之比。為一未知數，通常以可能誤差與近似值之比代替之。

歷屆普考特考試題

1-1 解釋下列名詞之意義：

- ① 次級資料 (Secondary date) ；
- ② 比例抽樣 (Proportional sampling) ；
- ③ 連續數列 (Continous series) ；
- ④ 全距 (Range) ；
- ⑤ 季節指數 (Seasonal index numbers) (48 年普考) 。

解 ① 次級資料：原始資料經整理、分析、簡化所得之結果，以供利用之數據資料稱之。

② 比例抽樣：將研究之對象依某種標準分為若干類，再依各類之大小，於其中抽出等比例之抽樣單位，以組成樣本之抽樣方法稱之。

例：袋中有紅球 20 顆，白球 30 顆，黑球 50 顆，即 $N_A = 20$ ， $N_W = 30$ ， $N_B = 50$ ，今欲抽取 10 球。若依比例抽樣法抽之，則當抽出紅球 2 顆，白球 3 顆，黑球 5 顆。

$$\text{即 } \frac{n_A}{N_A} = \frac{2}{20} = \frac{1}{10}$$

$$\frac{n_W}{N_W} = \frac{3}{30} = \frac{1}{10}$$

$$\frac{n_B}{N_B} = \frac{5}{50} = \frac{1}{10}$$

$$\therefore \frac{n_A}{N_A} = \frac{n_W}{N_W} = \frac{n_B}{N_B} = \frac{n}{N} = \frac{1}{10}$$

$$\begin{aligned} \text{其中 } N &= N_A + N_W + N_B \\ &= 20 + 30 + 50 = 100 \\ n &= n_A + n_W + n_B \\ &= 2 + 3 + 5 = 10 \end{aligned}$$

$$\frac{n}{N} = \frac{10}{100} = \frac{1}{10}$$

- ③連續數列：凡變量之任何兩數值間，可視為無限之數值介於其間者。
- ④全距：一群數據中，最大一數與最小一數之差稱之，亦稱兩極差。
- ⑤季節指數：乃季節變動指數之簡稱，表示以一年為週期之波動大小。

1 - 2 解釋下列各名詞，並說明其相互不同之點：

- ①連續數列與非連續數列。
- ②單對數圖與算術尺度圖。
- ③定基指數與連環指數 (Chain index number) (50 年普考)

解 ①連續數列：凡變量之兩任何數值間，可視為無限之數值介於其間者。

非連續數列：凡兩變量數值間只存在有限個不同之數值者。
連續數列與非連續數列間最大不同之點，在於任意兩變量數值間是否可視為存在無限數值，且連續數值係一種人為性的數列，而非連續數列是一種自然性的數列。

- ②單對數圖：任取縱橫兩軸之一為算術尺度，另一軸為對數尺度之圖形稱之。對數尺度是根據各數值之對數而劃定。
算術尺度圖：亦稱算術圖，係指縱橫兩軸均為算術尺度之圖形，其在算術尺度上之分點是等距離的。

兩者之不同點：

- (1)縱橫兩軸中，其中一軸所用之尺度不同。
- (2)表示變量各數值之尺度上，單對數尺度等距離表示等比，而算術尺度等距離代表等差。
- (3)因 $\log_{10} 1 = 0$ ，故單對數因沒零線，然而算術尺度圖有零線。
- ③定基指數：係採用固定基準的指數稱之。

連環指數：係以環比之各環相乘所得之指數稱之，亦稱鎖比指數。

兩者不同之點在於計算指數的基準不同。

1-3 解釋下列名詞：

- ①抽樣 (Sampling) ②樣本 (Sample)
 ③變數 (Variable) ④變值 (Variate value)
 ⑤統計單位 (Statistitical unit)。(51 年普考)

- 解** ①抽樣：自母群體中抽取部份單元，加以調查以作為分析研究稱之。
 ②樣本：藉抽樣而得之統計資料稱之。
 ③變數：統計單位可量的特性稱之。
 ④變值：凡一變數均具有許多的不同數量，此數量稱之為變值。如年齡為變數，則 30 歲、55 歲、87 歲均為變質。
 ⑤統計單位：乃為某特定之個體，為調查統計資料之對象。

1-4 解釋下列名詞：

- ①變數， ②變值， ③定基指數，
 ④連環指數， ⑤隨機抽樣。(55 年普考)

- 解** ①、②見第 1-3 之③④。
 ③、④見第 1-2 之③。
 ⑤隨機抽樣：於抽取母群體之個體以組成樣本時，不作任何人為因素及目的之選擇，純粹偶然地抽出個體，使每一個體皆具有相同被抽出之機率。

1-5 抽樣調查之目的為何？常用之抽樣方法有那幾種？試說明之。
 (57 年丙等特考、57 年普考：統計人員)

- 解** ①抽樣調查之目的在於省時、省事、省錢之原則下，以最少的代價，獲得較精確的部份資料，然後運用機率理論，以推求母群體的特徵值。

②常用之抽樣方法如下：

- (1)立意抽樣：吾人依據某種標準，判斷母群體中某些個體足具代表性，因而將其抽出組成樣本，此種抽樣方法稱之。
 此法適用於調查結果須具有典型代表性之調查計劃中。

- (2)隨機抽樣法：見 1-4 之⑤。

- (3)分層抽樣法：將母體依某種標準劃分為若干層次 (stratum)，並使層內之抽樣單位盡量等質，層間之抽樣單位盡量

異質，以提高樣本之精密度。再由各層任抽出適當之單位組成一組樣本，並根據此項樣本以推論母體之狀況之方法稱之。

(4)集體抽樣法：（或稱部落抽樣法）將母體分為若干集體（cluster），並使集體內之抽樣單位儘量異質，集體間之抽樣單位儘量同質，再定此等集體為抽出單位，依單純任抽方法抽出其中若干集體，在此抽出之若干集體內舉行調查其全體，最後根據此項資料以推論其母體之方法稱之。

(5)多段抽樣法，自母群體抽出個體以組樣本時，係分成若干段來進行之抽樣方法稱之。當前段各組間差異不大，後段各組內所含單位甚多時採用。

此外，尚有系統抽樣法、牛曼分層抽樣法、兩段抽樣法等多種。各抽樣方法之性質互異，抽樣調查時，宜根據調查對象之性質及其所處環境之不同，選用最適當的方法，以使得在相同經費下，獲得最高的抽樣效果。

1-6 述明普查與抽樣調查之目的，並指出其異同所在，在統計應用上，兩種調查之評價為何？

解 ①普查之目的在於獲得完全詳實可靠的資料，直接求算母群體之特徵值，抽樣調查之目的，則在省時、省錢、省事的原則下，以最少的代價，獲得較精確的部份資料，然後運用機率原理，以推求母體的特徵值。

②普查與抽樣調查均為統計調查，目的均在獲取母群體的特徵值，為統計方法上不可或缺之調查，故其方案設計及調查執行之作業過程，除有關調查方式的部份外，其餘大致相同。普查與抽查主要之不同點如下：

(1)空間方面：普查係普遍性的全面調查，抽查則僅調查群體中的一部份。

(2)時間方面：普查最重視其一致性，以免重複或遺漏，且辦理期限較久，抽樣雖對靜態資料亦要求時間一致；但部份場合則不必強求時間一致，且作業期間較短。例如品質管制方面。

(3)方法方面：普查講究方法統一；抽查可綜合應用各種方法

，且抽查須應用機率原理，抉擇抽樣技術方法。

(4)問項方面：普查之問項簡易；抽查則在必要時得更繁複深入，故問項特別廣泛。

(5)組織方面：普查動員多，須有健全組織，以靈活指揮，便利管理；抽查規模較小，可免專設機構而各機關兼辦。

(6)經費方面：普查經費浩繁，僅政府有能力辦理；抽查花費低廉，私人亦能辦理。

(7)確度方面：普查可消除抽樣誤差，而含有較大之非抽樣誤差，但非抽樣誤差可減低甚多。

- ③兩種調查之評價：抽查為普查之一種替換使用的調查方法，二者各有其優劣點：就統計方法上而言，抽查可能存有某種程度之抽樣誤差，卻可相對降低其費用，且在某些情況下，抽查因非抽樣誤差之巨幅縮小，致總誤差小於普查之誤差，而較普查更具精確性，且抽查又可適用於不便辦理普查之試驗性調查，故一般而言，抽查有利於普查，且有代替普查之勢。但抽查所得之長期時間數列又須普查資料予以驗正，因而普查仍具其重要性。

1-7 原始資料之分類標準與原則為何？（57年特考、58年升等考試）

解 ①分類標準：

- (1)時間標準：以事物發生之時間先後而加以分類。
- (2)地域標準：以事物發生之區域而加以分類。
- (3)特性標準：依事物之質特性或量特性加以分類。

②分類原則：

- (1)互斥性：所分之各類，必須互相排斥，凡能歸入某類者，絕不能再歸入他類，此目的在於分類時避免重複。
- (2)周延性：所分之各類必須能包括所應分類的一切事物。

1-8 統計資料整理的方法有幾種？試說明其步驟並比較其優劣點。（48年普考）

解 ①劃記法：先依照已定之分類項目，做一分類整理表，表中標明所分之類別。再根據原始調查表或登記表之資料，依次用記號劃於表上相等之欄內，最後總記各欄內之記號總數填入

合計欄，則歸類工作即告完成。

- ② 卡片法：先將原始資料分錄於大小相同之卡片上，再依所定之類別，依序分放卡片，最後點計各類卡片張數，即為各組之次數。
- ③ 電腦分類法：乃將資料之特性編成代號打在電腦卡上，然後送入電腦，則它將所有的打孔卡片依各類別而分開，完成歸類工作。

上述三種方法中，以劃記法最簡便，但校對困難，一旦有錯誤則不易發覺；卡片法因須分錄卡片，故手續繁冗，但查對容易；電腦分類法迅速正確，惟費用較高，僅宜用於大量分類。

1-9 解答下列各問題：

- ① 統計結果，能否百分之百正確？ ③ 何謂統計單位？
- ② 單靠統計學的知識，能否做好各種統計工作？
- ④ 物價指數公式有所謂型偏誤。何謂型偏誤？怎樣判斷公式有

無型偏誤？（56年普考）

解 ① 統計資料係一群數值，可分為二類：一為精確數值，不存在任何誤差；一為近似數值，含有若干誤差之數值。統計資料大半為近似數值，且大都依抽樣調查而得，故依統計而得之結果，不能百分之百的正確。

② 統計的最後且最重要之工作，在於統計結果之解釋，惟解釋統計分析之結果，頗非易事。必須對於統計方法之性能及其應用上之限制，有著明確的認識；對於研究對象有關之科學或業務，有著相當的瞭解，然後運用正確之判斷能力，審慎予以解釋，才可免於舛謬。故單靠統計學的知識，未必能做好統計工作。

③ 統計單位見第1-3之⑤。

④ 物價指數之前進指數與後退指數的乘積等於一時，即合於時間互換測驗。簡單指數公式不合於時間互換測驗者，則謂為型偏誤，因其係由於計算指數公式型式而來之定向誤差，故稱之。判斷指數公式有無型偏誤，視其前進指數與後退指數之相乘積是否等於一而定。若等於一，則為無型偏誤；若大於一，則為型偏高；若小於一，則為型偏低。

1 - 10 解釋下列名詞之意義：

- ①時間數列。
- ②組限。
- ③環比。
- ④長期趨勢。
- ⑤正相關。 (48年普考)

- 解 ①時間數列：一數量變動依時間先後順序安排之觀測值所成的集合稱之。
- ②組限：於分組次數表中，每一組均有兩個限值，數值較小者稱下限，數值較大者稱上限。此二上下限即為組限，其決定分組次數表中一組之範圍。
- ③環比：指數之基期依次移動者稱之；可表示每兩相鄰兩時期物價的相對變動。
- ④長期趨勢：乃一反映繼續上升、停滯或下降之圓滑與規律變動之時間數列。
- ⑤正相關：兩變量相對應之數值有同時增加及同時減少之一致趨勢稱之。

1 - 11 試述繪製統計圖之步驟。(50年普考，59年特考)。

- 解 ①甄選資料：依所欲研究之目的，善選所切用之資料。
- ②選擇圖式：依資料的性質及繪圖之目的，選擇最適當的圖式。其選擇標準約略如下：
- (1)欲表現某現象在時間上之變動情形，則宜採用歷史線圖；亦可採用綜長條圖，但不如前者適宜。
 - (2)欲表現某現象在各地區之分佈狀況，宜用統計地圖。
 - (3)欲表現某現象之次數分配狀況，宜用次數多邊圖或直方長條圖；若為非連續變量的次數資料，亦可用縱長條圖形表之。
 - (4)欲表現各部分細數與總數之比例關係，宜用圓形圖。
 - (5)欲表現某現象在時間之相對變動，宜用單對數尺度之歷史線圖；如欲表現絕對變動，宜用算術尺度之歷史線圖。
 - (6)欲表現以上或以下之累積次數之變動情形，宜用累積次數多邊圖或直方圖。

- (7) 屬性數列之資料宜用長條圖表示。
- (8) 宣傳或廣告用之圖表，宜採用鮮明的色彩或繪形象圖。
- ③ 分析計算：若原來數據不能直接應用，須化為百分數等時，須經過分析計算的手續。
- ④ 確定圖形輪廓：視資料的性質及圖形之種類，確定適當寬度、長度，不宜使圖形過於扁平或過於狹長。至於輪廓線是否繪出，並無硬性地規定，可任意變動。
- ⑤ 繪基線、定比度：基線示稱主軸，凡採用坐標顯示的圖形，於圖廓確定後，即著手安置縱橫基線並按分類之多少、數量情形決定比度，但須兼顧明確清楚、美觀悅目。
- ⑥ 定圖例、繪草圖：於圖中若有數種資料同時表現，為區別辨識，可用不同線條或顏色表示；但應設定圖例說明，以便觀察了解，並隨即以鉛筆依統計資料、比例等繪製草圖。
- ⑦ 審核草圖：當草圖繪製完畢，應加以審核，端慮其形式大小是否適當？線圖位置是否正確？資料有無遺漏或錯誤？若發現任何不當，即加以修正及補充。
- ⑧ 上墨著色：草圖經審核無誤後，即可上墨著色，繪製正圖；著色時應注意線圖之粗細、墨色之濃淡，顏色之調和等事項。
- ⑨ 清潔圖面：墨色全乾後，將圖面上不潔之污漬筆痕予以去除，力求圖面之清潔美觀。
- ⑩ 校對全圖：為慎重起見，於圖全部繪製完成後，應將圖中各部份再仔細核對一次，以求圖形之精確。

1 - 12 我們於同一圖中，作兩條歷史曲線，常需配合縱尺度，何故？配合的方法有那幾種？（56年普考）

解 ① 配合縱尺度，在於研究經濟問題或其他問題時，時或需在同一圖中作兩條曲線，代表兩種事實，以便觀察二者之間有無關連。此二種事實數量的單位或不相同，或相同而數量之差異很大，於是需要配合兩個縱尺度，其目的在使：

- (1) 兩條線相去不遠而便於比較。
- (2) 兩條線的起伏適合事實。

② 配合縱尺度的方法有四種：

- (1) 將原有兩種事實求得平均數，然後以其平均數之比，或其